UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/660,325	09/10/2003	Dinei A. Florencio	MCS-030-03 (303705.01)	5954

27662   7590   05/08/2007
MICROSOFT CORPORATION
C/O LYON & HARR, LLP
300 ESPLANADE DRIVE
SUITE 800
OXNARD, CA 93036

EXAMINER
LERNER, MARTIN

ART UNIT	PAPER NUMBER
2626	

MAIL DATE	DELIVERY MODE
05/08/2007	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

PTOL-90A (Rev. 04/07)

-- *The MAILING DATE of this communication appears on the cover sheet with the correspondence address* --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) ☐ Responsive to communication(s) filed on _____.

2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.

3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) ☒ Claim(s) *1 to 36* is/are pending in the application.

    4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5) ☐ Claim(s) _____ is/are allowed.

6) ☒ Claim(s) *1 to 12, 18 to 20, 22 to 24, 26 to 29, and 31 to 36* is/are rejected.

7) ☒ Claim(s) *13 to 17, 21, 25, and 30* is/are objected to.

8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) ☒ The specification is objected to by the Examiner.

10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.

    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

    Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

    a) ☐ All   b) ☐ Some * c) ☐ None of:

      1. ☐ Certified copies of the priority documents have been received.

      2. ☐ Certified copies of the priority documents have been received in Application No. _____.

      3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

    * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☒ Information Disclosure Statement(s) (PTO/SB/08)
    Paper No(s)/Mail Date _____.

4) ☐ Interview Summary (PTO-413)
    Paper No(s)/Mail Date. _____.

5) ☐ Notice of Informal Patent Application

6) ☐ Other: _____.

DETAILED ACTION

Specification

1.   The disclosure is objected to because of the following informalities:

On page 8, line 2, "in" should be inserted after "used".

On page 17, line 30, "and" should be deleted after "as well as".

On page 19, line 14, "many time" should be "many times".

On page 38, line 26, "invoiced" should be "unvoiced".

On page 42, line 8, "widow" should be "window".

On page 44, line 29, "widow" should be "window".

On pages 46 to 48, Steps 700, 710, and 760 are illustrated in Figure 7, but are not described in the Specification.

Appropriate correction is required.

Claim Rejections - 35 USC § 102

2.   The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

3.   Claims 1 to 6, 26, 27, 29, and 32 are rejected under 35 U.S.C. 102(b) as being anticipated by *Leitch et al.*

Regarding independent claims 1 and 26, *Leitch et al.* discloses a system and computer-implemented process, comprising:

"receive one or more sequential frames of a digital audio signal" – an input speech signal of a voice message is digitized and stored (column 19, line 67 to column 20, line 3); input speech is framed into 20 ms blocks (column 24, lines 10 to 12);

"decode each frame of the digital audio signal as it is received" – digitized voiced samples can be stored in formats including LPC based forms (column 16, lines 20 to 22); implicitly, a speech signal encoded by LPC (linear predictive coding) must be decoded when it is received; WSOLA-SD requires a determination of a pitch period of voiced portions (column 22, lines 21 to 24: Figure 25); obtaining a pitch period of a speech signal involves decoding the speech signal;

"determine a content type of segments of the decoded audio signal from a group of predefined segment content types, each segment content type having an associated type-specific temporal modification process" – an energy per block and zero-crossing rate are computed, and an energy threshold is determined to detect voiced speech as a function of energy per block; using an energy threshold and a zero-crossing threshold, contiguous blocks of voiced speech of length of at least 5 blocks are located, and pitch analysis is performed on all voiced segments; segments that are not marked as voiced speech are now marked as tentative unvoiced segments; contiguous blocks of at least 5 frames in the 'tentative unvoiced segments' are taken and pitch analysis is done (column 24, lines 13 to 35); thus, segments are marked as voiced or unvoiced ("determine a content type of segments. . . from a group of predefined segment content

types"); then, time-scaling is done in accordance with whether the segment is voiced or unvoiced, where a voiced segment is compressed with a segment size Ss = 2*Pitch, and an unvoiced segment is compressed with a segment size Ss = 100 (column 24, lines 35 to 53); compression of audio by a method where a voiced or unvoiced speech signal has a different segment size is "an associated type-specific temporal modification process";

"modify a temporal scale of one or more segments of the decoded audio signal using the associated type-specific temporal modification process specific to each segment content type" – WSOLA-SD provides time-scaling in accordance with whether the segment is voiced or unvoiced, where a voiced segment is compressed with a segment size Ss = 2*Pitch, and an unvoiced segment is compressed with a segment size Ss = 100 (column 24, lines 35 to 53).

Regarding claims 2 and 3, *Leitch et al.* discloses that, initially, a classification of a frame as voiced or unvoiced is based only on an energy threshold, zero-crossing threshold, and pitch analysis for 20 ms blocks ("based solely on the frame being classified"); subsequently, 'tentative unvoiced segments' are taken and a pitch analysis is done on contiguous blocks of at least 5 frames to determine whether the segment is voiced or unvoiced ("at least partially based on information derived from one or more neighboring frames") (column 24, lines 16 to 35).

Regarding claim 4, *Leitch et al.* discloses determining whether a speech segment is voiced or unvoiced for input speech of 20 ms blocks (column 24, lines 1 to 12); implicitly, blocks of input speech are expanded or compressed sequentially.

Regarding claim 5, *Leitch et al.* discloses that the analysis of segment size Ss is dependent on the pitch period of the input speech for time-scale expansion or compression using WSOLA-SD (column 23, lines 46 to 56); a pitch period is equivalent to "a periodicity of each data frame."

Regarding claims 6 and 27, *Leitch et al.* discloses determining whether the speech segment is voiced or unvoiced for time-scale expansion or time-scale compression by WSOLA-SD (column 24, lines 1 to 9).

Regarding claims 29 and 32, *Leitch et al.* discloses that the WSOLA algorithm provides for time-scaling expansion if the time-scaling parameter α less than 1, and time-scaling compression if α is greater than 1 (column 19, lines 59 to 63); thus, α represents "a target temporal modification ratio"; for time scale expansion ("stretching"), samples are copied and added (column 20, lines 3 to 65); expansion of a segment is in accordance with adding segments that are the size of a pitch period, as Ss = 2*Pitch for voiced speech ("by approximately one or more pitch periods to increase a length of the at least one voiced type segment") (column 24, lines 1 to 8).

Claim Rejections - 35 USC § 103

4.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

5.   Claims 7 to 9, 11, 20, 22, 24, and 28 are rejected under 35 U.S.C. 103(a) as

being unpatentable over *Leitch et al.* in view of *Ananthpadmanabhan et al.*

Concerning independent claim 8, *Leitch et al.* omits only determining a content

type for, and temporally modifying, "mixed segments". However, it is generally well

known that speech signals may be classified into more than simply voiced and unvoiced

segments, and that many methods of classifying speech segments involve

considerations of segments that have properties that are intermediate between voiced

and unvoiced segments. *Ananthpadmanabhan et al.* teaches a method of speech

coding, where a mode classification module 408 selects a particular encoding mode 410

for a current frame based upon the periodicity of the frame, as voiced, unvoiced, or

transient. Transient frames are typically transitions between voiced and unvoiced

speech. (Column 7, Line 59 to Column 8, Line 18: Figures 5 and 6) An objective is to

achieve a significant data rate reduction by the use of speech analysis, and applying

different coding-decoding algorithms to different types of speech frames. (Column 1,

Lines 23 to 26; Column 2, Lines 22 to 43) It would have been obvious to one of

ordinary skill in the art to compress a speech signal by mode specific encoding of mixed

segments as suggested by *Ananthpadmanabhan et al.* in a voice compression method

of *Leitch et al.* for a purpose of achieving a significant data rate reduction by the use of

speech analysis.

Concerning claims 7, 20, and 28, similar considerations apply.

Concerning claim 9, *Leitch et al.* discloses a voiced segment is compressed with

a segment size Ss = 2*Pitch, which is approximately the size of one pitch period in

length, after an overlap of Ss/2 samples (column 20, lines 10 to 13; column 24, lines 1

to 8).

Concerning claims 11, 22, and 24, *Leitch et al.* discloses that the WSOLA

algorithm provides for time-scaling expansion if the time-scaling parameter α less than

1, and time-scaling compression if α is greater than 1 (column 19, lines 59 to 63); thus,

α represents "a target temporal modification ratio"; for time scale expansion

("stretching"), samples are copied and added (column 20, lines 3 to 65); expansion of a

segment is in accordance with adding segments (column 24, lines 1 to 8).

6.   Claims 10, 12, 23, 31, and 33 are rejected under 35 U.S.C. 103(a) as being

unpatentable over *Leitch et al.* in view of *Ananthpadmanabhan et al.* as applied to

claims 8, 11, 22, 26, and 27 above, and further in view of *Yeldener.*

Concerning claims 10 and 31, *Leitch et al.* discloses computing a normalized

cross-correlation for time compression or time expansion to find the best correlating

samples, but does not expressly say that a maximum peak is compared to thresholds

for determining a content type. However, *Yeldener* teaches finding a pitch by

maximizing a normalized cross-correlation function of peak amplitudes, and comparing

a probability parameter Pv with a pre-specified threshold to determine whether a

previous frame was voiced. (Column 11, Lines 6 to 35; Column 12, Lines 32 to 49) An

advantage is that once a voiced signal is established, its pitch varies only within a

limited range, reducing the probability of encountering a pitch doubling problem.

(Column 11, Lines 21 to 38) It would have been obvious to one having ordinary skill in

the art to compare a maximum peak to a threshold for determining a content type of speech as suggested by *Yeldener* in a voice compression method of *Leitch et al.* for a purpose of reducing a pitch doubling problem in tracking a pitch.

Concerning claims 12, 23, and 33, *Leitch et al.* discloses a method of obtaining a best matching waveform for copying by a normalized cross-correlation function ("identifying at least one of the segments as a template"), and aligning and merging the matching segments by a ramp function for expanding ("stretching") a voice signal (column 20, lines 20 to 47); correspondingly, segments are "cut out", or deleted, for compression.

7.　　Claims 18, 34, and 36 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Leitch et al.* in view of *Ananthpadmanabhan et al.* as applied to claims 8 and 26 above, and further in view of *Mozer.*

Generally, *Leitch et al.* discloses expansion of a speech signal by adding segments including unvoiced segments having a segment size Ss = 100 (column 20, lines 15 to 47; column 24, lines 41 to 53), but does not, specifically, say how the process is performed for unvoiced segments, nor expressly determine an insertion point for unvoiced segments. However, *Mozer* teaches time domain compression and synthesis for unvoiced audible signals, where unvoiced sounds, beginning and ending at quasi random points for the duration of any desired interval, are repeatedly reproduced during synthesis a sufficient number of times to reconstruct a time segment. An objective is to eliminate a characteristic buzz, or a noticeable periodicity, of repeated

segments. (Column 3, Line 56 to Column 4, Line 32) It would have been obvious to one having ordinary skill in the art to utilize the method of repeating unvoiced segments as taught by *Mozer* in a voice compression method of *Leitch et al.* for a purpose of eliminating a characteristic buzz or noticeable periodicity.

8.   Claims 19 and 35 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Leitch et al.* in view of *Ananthpadmanabhan et al.* and *Mozer* as applied to claims 8, 18, 26, and 34 above, and further in view of *Macon et al.* *("Sinusoidal Modeling and Modification of Unvoiced Speech")*.

Leitch et al. omits the concept of randomizing the phase of a synthetic unvoiced segment by Fourier transforming, introducing a random rotation of the phase, and inverse Fourier transforming a segment. However, *Macon et al.* teaches time scale modification of unvoiced speech, where a phase randomization procedure is performed in a frequency domain, and an inverse Fourier transform is performed before an overlap-add in analysis-by-synthesis/overlap-add (ABS/OLA). An objective is to eliminate a tonal artifact. (Pages 558 to 559) It would have been obvious to one having ordinary skill in the art to perform a phase randomization of an unvoiced segment as taught by *Macon et al.* in a voice compression method of *Leitch et al.* for a purpose of eliminating a tonal artifact.

Allowable Subject Matter

9.   Claims 13 to 17, 21, 25, and 30 are objected to as being dependent upon a

rejected base claim, but would be allowable if rewritten in independent form including all

of the limitations of the base claim and any intervening claims.

Conclusion

10.   The prior art made of record and not relied upon is considered pertinent to

Applicants' disclosure.

Bhadkamkar et al. and Kleijn disclose related art.

Any inquiry concerning this communication or earlier communications from the

examiner should be directed to Martin Lerner whose telephone number is (571) 272-

7608. The examiner can normally be reached on 8:30 AM to 6:00 PM Monday to
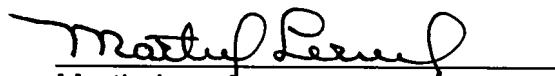
Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, David R. Hudspeth can be reached on (571) 272-7843. The fax phone

number for the organization where this application or proceeding is assigned is 571-

273-8300.

Information regarding the status of an application may be obtained from the

Patent Application Information Retrieval (PAIR) system. Status information for

published applications may be obtained from either Private PAIR or Public PAIR.

Status information for unpublished applications is available through Private PAIR only.

For more information about the PAIR system, see http://pair-direct.uspto.gov. Should

you have questions on access to the Private PAIR system, contact the Electronic

Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a

USPTO Customer Service Representative or access to the automated information

system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

ML
5/2/07

Martin Lerner
Examiner
Group Art Unit 2626